

NONPARAMETRIC ESTIMATION OF THE JUMP RATE FOR PIECEWISE-DETERMINISTIC MARKOV PROCESSES

ROMAIN AZAÏS, FRANÇOIS DUFOUR, AND ANNE GÉGOUT-PETIT

ABSTRACT. This paper presents a nonparametric method for estimating the conditional density associated to the jump rate of a piecewise-deterministic Markov process. In our framework, the estimation needs only one observation of the process within a long time interval. Our method relies on a generalization of Aalen's multiplicative intensity model. We prove the uniform consistency of our estimator, under some reasonable assumptions related to the primitive characteristics of the process. A simulation example illustrates the behavior of our estimator.

1. INTRODUCTION

This paper is devoted to the nonparametric estimation of the jump rate for piecewise-deterministic Markov processes, from only one observation of the process within a long time interval. Our approach is based on methods investigated in the previous work of the authors [6] and a generalization of the well-known multiplicative intensity model, developed by Aalen in [1, 2, 3] in the middle of the seventies.

Piecewise-deterministic Markov processes (PDMP's) have been introduced in the literature by Davis in [13] as a general class of non-diffusion stochastic models. They are a family of Markov processes involving deterministic motion punctuated by random jumps. The path depends on three local characteristics namely the flow Φ , the jump rate λ , which determines the interarrival times, and the transition kernel Q , which specifies the post-jump location. A suitable choice of the state space and the local characteristics Φ , λ and Q provides stochastic models covering a large number of problems, for example in reliability (see [13] and [8]). Denote by f the conditional density of the interarrival times associated to λ . This is a function of two variables: a spatial mark and time. The purpose of this paper is to develop a nonparametric procedure to estimate this function, when only one observation of the process within a long time is available. To the best of our knowledge, the estimation of the conditional distribution of the interarrival times for this class of stochastic models has never been studied. Furthermore, this paper relies on [6] in which we focus on the nonparametric estimation of the jump rate and the cumulative rate for a class of non-homogeneous marked renewal processes. This class of stochastic models amounts to considering a particular piecewise-deterministic process, whose post-jump locations do not depend on interarrival times.

2010 *Mathematics Subject Classification.* Primary: 62M05, Secondary: 60J25.

Key words and phrases. Piecewise-deterministic Markov process, ergodicity of Markov chains, nonparametric estimation, jump rate estimation, Nelson-Aalen estimator, asymptotic consistency.

This work was supported by ARPEGE program of the French National Agency of Research (ANR), project "FAUTOCOES", number ANR-09-SEGI-004.

As counting processes may model a large variety of problems, nonparametric and semi-parametric estimation methods have been developed by many authors for their statistical inference. The famous multiplicative intensity model has been extensively investigated by Aalen since 1975 (see [1, 2, 3]). This model postulates the existence of a predictable process Y and a deterministic function λ , called the jump rate or the hazard rate, such that the stochastic intensity of the underlying counting process N is given by the product $Y\lambda$. In this context, Aalen provided a useful method for estimating the cumulative rate $\Lambda(t) = \int_0^t \lambda(s)ds$. The associated consistent estimator is now called the Nelson-Aalen estimator. In 1983, Ramlau-Hansen focused on smoothing this estimator by some kernel methods, in order to estimate directly the jump rate λ . A nonparametric estimate of the rate λ has been provided in [23].

As mentioned before, a large number of estimation problems are related to the estimation of jump rates depending on both time and a spatial variable. The Nelson-Aalen estimator is proved to be well-adapted for a large variety of developments and applications (see the book [5] and the references therein), in particular in survival analysis, or in statistics of processes. For example, one may apply Aalen's approach for estimating the jump rate of a marked counting process, whose state space is finite, from independent observations. More recently, Comte *et al.* proposed in [10] a new strategy for the inference for counting processes in presence of covariates, under the multiplicative assumption.

Semiparametric estimation methods have been mainly investigated in presence of continuous covariates, beginning with Cox [11]. One may refer the reader to the book [5] and the references therein for a large review of the literature on these models. There exists also an extensive literature on nonparametric approaches when the spatial mark takes its values in a continuous space. We do not attempt to present an exhaustive survey on this topic, but refer the interested reader to [4, 5, 16, 19] and the references therein for detailed expositions on these techniques. In particular, McKeague and Utikal proposed in 1990 a nonparametric estimator of the jump rate when the covariate belongs to $[0, 1]$ (see [20]). Their approach is based on a generalization of Aalen's multiplicative model. It consists in smoothing a Nelson-Aalen type estimator both in spatial and time directions. The authors demonstrated the uniform convergence in probability of their estimator. Li and Doss extended in turn McKeague and Utikal's work for the multidimensional case in [18]. This paper relies on a local linear fit in the spatial direction. Their theoretical results concern the weak convergence of the proposed estimators. The interested reader may also consult the papers written by Utikal [25, 26]. These two papers deal with the nonparametric estimation of the jump rate for two special classes of marked counting processes, observed within a long time, under some continuous-time martingale assumptions. The Euclidean structure of the covariate state space plays a key role in the papers mentioned above. At the same time, the nonparametric approach has been considered by Beran [7], Stute [24] and Dabrowska [12], but in the independent and identically distributed case.

Our approach relies on our previous work [6] and a generalization of Aalen's multiplicative model. The main difficulty is related to the dependence of the transition kernel on the previous interarrival time. This excludes the techniques developed in the literature [6, 18, 20, 25, 26] for estimating the jump rate λ . The keystone of the present paper is to consider the Markov chain $(Z_n, Z_{n+1}, S_{n+1})_{n \geq 0}$, where the Z_n 's denote the post-jump locations of the process, and the S_n 's denote the interarrival times. The main idea in this work is to deal with the conditional distribution of S_{n+1} given Z_n and Z_{n+1} . We prove that

this conditional distribution admits a jump rate $\tilde{\lambda}$, under some regularity conditions on the primitive data of the process (see Proposition 3.8). Furthermore, a conditional independence result is satisfied by the discrete-time process $(Z_n, Z_{n+1}, S_{n+1})_{n \geq 0}$ (see Proposition 5.8). In this context, we focus on the estimation of the jump rate $\tilde{\lambda}$, which is a function of three variables: two spatial marks and time. Moreover, the two spatial variables take their values on a general metric space. In particular, this rules out the procedures investigated by the authors mentioned above [18, 20, 25, 26]. As a consequence, our method consists in involving a thin partition of the state space. We take a leaf out of a few proofs of our previous work [6] for estimating the function $\tilde{l}(A, B, t)$, which is an approximation of the jump rate $\tilde{\lambda}(x, y, t)$, for $x \in A$ and $y \in B$. In the rest of the paper, we use the convergence property of this estimator to tackle the estimation problem of the density of interest f . We study the connection between $\tilde{\lambda}$ and the conditional density f (see Proposition 6.7). This link involves the conditional probability $\mathbf{P}_\nu(S_1 > t, Z_1 \in B | Z_0 \in A)$, where ν denotes the invariant measure of the process $(Z_n)_{n \geq 0}$. An efficient estimate of this quantity is presented in Proposition 6.10. We finally provide a nonparametric estimator of the conditional density f , and we prove a result of uniform convergence in Theorem 6.11. An inherent difficulty is related to the presence of constrained jumps, when the process reaches the boundary of the state space. Moreover, we ensure the consistency of our estimator, under interpretable and reasonable conditions related to the primitive characteristics of the process.

The paper is organized in the following way. We first recall, in Section 2, the definition of a PDMP. Next, we introduce a new process, whose transition kernel does not depend on time. Section 3 is devoted to the existence of the conditional jump rate $\tilde{\lambda}$ of this Markov process (see Proposition 3.8). In Section 4, we state some ergodicity results, which will be useful in order to prove the uniform convergence of our estimator. Section 5 relies on some proofs given in [6]. In this section, we propose an estimator of the function $\tilde{l}(A, B, t)$, which is an approximation of the jump rate $\tilde{\lambda}(x, y, t)$ for $x \in A$ and $y \in B$. The main contribution of the paper lies in Section 6. We provide a nonparametric estimator of the conditional density f associated to λ , by studying the link between f and the jump rate $\tilde{\lambda}$. Our main result of consistency is presented in Theorem 6.11. Finally, a numerical example is given in Section 7 for illustrating the good behavior of our estimator. The technical proof of a conditional independence result (Proposition 5.8) is deferred in Appendix A.

2. DEFINITION OF A PDMP

This section is devoted to the definition of a piecewise-deterministic Markov process on a separable metric space. The process evolves in an open subset E of a separable metric space (\mathcal{E}, d) . The motion is defined by the three local characteristics (λ, Q, Φ) .

- $\Phi : E \times \mathbf{R}_+ \rightarrow \overline{E}$ is the deterministic flow. It satisfies,

$$\forall \xi \in E, \forall s, t \geq 0, \Phi(\xi, t + s) = \Phi(\Phi(\xi, t), s).$$

For each $\xi \in E$, $t^*(\xi)$ denotes the deterministic exit time from E :

$$t^*(\xi) = \inf\{t > 0 : \Phi(\xi, t) \in \partial E\},$$

with the usual convention $\inf \emptyset = +\infty$.

- $\lambda : \bar{E} \rightarrow \mathbf{R}_+$ is the jump rate. It is a measurable function which satisfies,

$$\forall \xi \in E, \exists \varepsilon > 0, \int_0^\varepsilon \lambda(\Phi(\xi, s)) ds < +\infty.$$

- Q is a Markov kernel on $(\bar{E}, \mathcal{B}(\bar{E}))$ which satisfies,

$$\forall \xi \in \bar{E}, Q(\xi, \bar{E} \setminus \{\xi\}) = 1 \quad \text{and} \quad Q(\xi, E) = 1.$$

There exists a filtered probability space $(\Omega, \mathcal{A}, (\mathcal{F}_t)_{t \geq 0}, \mathbf{P}_{\nu_0})$, on which the process $(X_t)_{t \geq 0}$ is defined (see [13]). The probability distribution of the initial value X_0 is ν_0 . Starting from $x \in E$, the motion can be described as follows. T_1 is a positive random variable whose survival function is,

$$\forall t \geq 0, \mathbf{P}_{\nu_0}(T_1 > t | X_0 = x) = \exp\left(-\int_0^t \lambda(\Phi(\xi, s)) ds\right) \mathbf{1}_{\{0 \leq t < t^*(x)\}}.$$

One chooses an E -valued random variable Z_1 according to the distribution $Q(\Phi(\xi, T_1), \cdot)$. Let us remark that the post-jump location depends on the interarrival time T_1 . The trajectory between the times 0 and T_1 is given by

$$X_t = \begin{cases} \Phi(\xi, t) & \text{for } 0 \leq t < T_1, \\ Z_1 & \text{for } t = T_1. \end{cases}$$

Now, starting from X_{T_1} , one selects the time $S_2 = T_2 - T_1$ and the post-jump location Z_2 in a similar way as before, and so on. This gives a strong Markov process with the T_k 's as the jump times (with $T_0 = 0$). One often considers the embedded Markov chain $(Z_n, S_n)_{n \geq 0}$ associated to the process $(X_t)_{t \geq 0}$ with $Z_n = X_{T_n}$, $S_n = T_n - T_{n-1}$ and $S_0 = 0$, that is, the Z_n 's denote the post-jump locations of the process, and the S_n 's denote the interarrival times.

On the strength of [13] (Chapter 1, Section 24 Definition of the PDP), the embedded chain $(Z_n, S_n)_{n \geq 0}$ is generated by a stochastic dynamic system. There exist two measurable functions φ and ψ , and two independent random sequences $(\varepsilon_n)_{n \geq 0}$ and $(\delta_n)_{n \geq 0}$, such that, for any $n \geq 1$,

$$(1) \quad \begin{cases} S_n &= \varphi(Z_{n-1}, \delta_{n-1}), \\ Z_n &= \psi(Z_{n-1}, S_n, \varepsilon_{n-1}). \end{cases}$$

It appears natural to introduce the one-jump counting process N^{i+1} , for each integer i , given for any $t \geq 0$ by

$$N^{i+1}(t) = \mathbf{1}_{\{S_{i+1} \leq t\}},$$

and $(\mathcal{F}_t^{i+1})_{t \geq 0}$ its natural filtration. Finally, for each integer n , denote by \mathcal{G}_n the σ -field $\sigma(Z_0, \dots, Z_n)$. For a matter of readability, we introduce also the following notations,

$$\begin{aligned} \forall \xi \in E, \quad \forall t \geq 0, \quad \bar{\lambda}(\xi, t) &= \lambda(\Phi(\xi, t)), \\ \forall \xi \in E, \quad \forall t \geq 0, \quad \forall A \in \mathcal{B}(E), \quad \bar{Q}(\xi, t, A) &= Q(\Phi(\xi, t), A). \end{aligned}$$

In all the sequel, let us denote by f and G the probability density function and the survival function associated to the jump rate $\bar{\lambda}$.

$$\begin{aligned} \forall \xi \in E, \quad \forall t \geq 0, \quad G(\xi, t) &= \exp\left(-\int_0^t \bar{\lambda}(\xi, s) ds\right), \\ \forall \xi \in E, \quad \forall t \geq 0, \quad f(\xi, t) &= \bar{\lambda}(\xi, t) G(\xi, t). \end{aligned}$$

The purpose of this paper is to provide a nonparametric estimate of the conditional density f . We shall prove the consistency of the estimator from some assumptions related to the characteristics of the process.

3. CONDITIONAL DISTRIBUTION OF S_{n+1} GIVEN Z_n, Z_{n+1}

In this section, we show that the conditional distribution of S_{n+1} given Z_n and Z_{n+1} admits a jump rate (see Proposition 3.8). This result directly induces a corollary about the compensator of the counting process N^{n+1} in a particular filtration (see Corollary 3.10). First, some assumptions on the characteristics of the process are presented. In particular, we assume that the transition kernel \bar{Q} admits a density according to an auxiliary measure μ on $(E, \mathcal{B}(E))$.

Assumption 3.1. *The transition kernel \bar{Q} may be written in the following way,*

$$\forall \xi \in E, \forall s \geq 0, \forall B \in \mathcal{B}(E), \bar{Q}(\xi, s, B) = \int_B \tilde{Q}(\xi, s, y) \mu(dy),$$

where μ is an auxiliary measure on $(E, \mathcal{B}(E))$.

Furthermore, we impose some conditions on both the measure μ and the associated density.

Assumptions 3.2.

- a) For any measurable set B with non-empty interior, $\mu(B) > 0$.
- b) For any $x, y \in E$, $\tilde{Q}(x, \cdot, y)$ is a continuous function.
- c) There exists $m > 0$ such that,

$$(2) \quad \forall x \in E, \forall s \geq 0, \forall y \in E, \tilde{Q}(x, s, y) \geq m.$$

Now, we impose some regularity conditions on the density f and the jump rate $\bar{\lambda}$.

Assumptions 3.3.

- a) $f(\xi, \cdot)$ is a continuous function for any $\xi \in E$.
- b) There exists a locally integrable function $M : \mathbf{R}_+ \rightarrow \mathbf{R}_+$ such that,

$$\forall \xi \in E, \forall t \geq 0, \bar{\lambda}(\xi, t) \leq M(t).$$

For the sake of clarity, let us introduce the following notations,

$$(3) \quad \forall \xi \in E, \forall s \geq 0, \forall y \in E, G\tilde{Q}(\xi, s, y) = G(\xi, s)\tilde{Q}(\xi, s, y),$$

$$(4) \quad \forall \xi \in E, \forall s \geq 0, \forall y \in E, f\tilde{Q}(\xi, s, y) = f(\xi, s)\tilde{Q}(\xi, s, y).$$

Furthermore, we denote also,

$$(5) \quad \forall \xi \in E, \forall s \geq 0, \forall B \in \mathcal{B}(E), G\bar{Q}(\xi, s, B) = G(\xi, s)\bar{Q}(\xi, s, B),$$

$$(6) \quad \forall \xi \in E, \forall s \geq 0, \forall B \in \mathcal{B}(E), f\bar{Q}(\xi, s, B) = f(\xi, s)\bar{Q}(\xi, s, B).$$

We shall establish that $G(x, t^*(x))$ and $G\tilde{Q}(x, t^*(x), y)$ are two strictly positive numbers for any $x, y \in E$.

Remark 3.4. Let $x, y \in E$. On the one hand, we have

$$(7) \quad G(x, t^*(x)) \geq \exp \left(- \int_0^{t^*(x)} M(s) ds \right) > 0,$$

because M is a locally integrable function. On the other hand, from equations (2) and (3), we have

$$(8) \quad G\tilde{Q}(x, t^*(x), y) \geq mG(x, t^*(x)) > 0.$$

Let us denote by R the transition kernel of the Markov chain $(Z_n)_{n \geq 0}$. In the following remark, we especially give an explicit formula for R .

Remark 3.5. The kernel R may be written in the following way [13, (34.12), page 116], for any $x \in E$, $B \in \mathcal{B}(E)$,

$$(9) \quad R(x, B) = \int_0^{t^*(x)} f\bar{Q}(x, s, B) ds + G\bar{Q}(x, t^*(x), B),$$

where the functions $G\bar{Q}$ and $f\bar{Q}$ have already been defined by (5) and (6). Furthermore, since the transition kernel \bar{Q} admits \tilde{Q} as a density according to the measure μ , one may also write

$$(10) \quad R(x, B) = \int_B \left[\int_0^{t^*(x)} f\tilde{Q}(x, s, y) ds + G\tilde{Q}(x, t^*(x), y) \right] \mu(dy).$$

Together with (7) and (8),

$$(11) \quad \begin{aligned} R(x, B) &\geq G(x, t^*(x)) \int_B \tilde{Q}(x, t^*(x), y) \mu(dy) \\ &\geq m\mu(B) \exp \left(- \int_0^{t^*(x)} M(s) ds \right). \end{aligned}$$

In particular, if $\overset{\circ}{B} \neq \emptyset$, $R(x, B) > 0$ because $\mu(B) > 0$ according to Assumptions 3.2.

Let us denote by ν_n (resp. by $\tilde{\nu}_n$, by η_n) the distribution of Z_n (resp. of the couple (Z_n, Z_{n+1}) , of (Z_n, Z_{n+1}, S_{n+1})). The following remark deals with the relation between $\tilde{\nu}_n$, ν_n and the transition kernel R .

Remark 3.6. Let n be an integer and $A \times B \in \mathcal{B}(E)^{\otimes 2}$. One may write

$$(12) \quad \begin{aligned} \tilde{\nu}_n(A \times B) &= \mathbf{P}_{\nu_0}(Z_{n+1} \in B, Z_n \in A) \\ &= \int_A R(x, B) \nu_n(dx). \end{aligned}$$

We focus our attention on the relation between the probability measures η_n and ν_n .

Remark 3.7. Let $t \geq 0$ and $B \in \mathcal{B}(E)$. We have

$$\begin{aligned} &\mathbf{P}_{\nu_0}(S_{n+1} > t, Z_{n+1} \in B | Z_n) \\ &= \mathbf{1}_{\{0 \leq t < t^*(Z_n)\}} \left[\int_{t \wedge t^*(Z_n)}^{t^*(Z_n)} f\bar{Q}(Z_n, s, B) ds + G\bar{Q}(Z_n, t^*(Z_n), B) \right]. \end{aligned}$$

This obviously induces that, for any $A, B \in \mathcal{B}(E)$ and $t \geq 0$,

$$(13) \quad \eta_n(A \times B \times]t, +\infty[) = \int_A \mathbf{1}_{\{0 \leq t < t^*(x)\}} \left[\int_{t \wedge t^*(x)}^{t^*(x)} f\bar{Q}(x, s, B) ds + G\bar{Q}(x, t^*(x), B) \right] \nu_n(dx).$$

The main result of this section lies on the following proposition. It deals with the existence of a jump rate for the conditional distribution of S_{n+1} given Z_n, Z_{n+1} .

Proposition 3.8. *Let n be an integer. The conditional distribution of S_{n+1} given Z_n, Z_{n+1} satisfies, for any $t \geq 0$,*

$$\mathbf{P}_{\nu_0}(S_{n+1} > t | Z_n, Z_{n+1}) = \exp \left(- \int_0^{t \wedge t^*(Z_n)} \tilde{\lambda}(Z_n, Z_{n+1}, s) ds \right) \mathbf{1}_{\{0 \leq t < t^*(Z_n)\}},$$

where the jump rate $\tilde{\lambda}$ is defined for any $x, y \in E$, $0 \leq t \leq t^*(x)$, by

$$(14) \quad \tilde{\lambda}(x, y, t) = \frac{f\tilde{Q}(x, t, y)}{\int_t^{t^*(x)} f\tilde{Q}(x, s, y) ds + G\tilde{Q}(x, t^*(x), y)}.$$

Proof. Let $x, y \in E$. $\tilde{\lambda}(x, y, \cdot)$ is a continuous function on the interval $[0, t^*(x)]$ because $f(x, \cdot)$ and $\tilde{Q}(x, \cdot, y)$ are two continuous functions in the light of Assumptions 3.2 and 3.3. The survival function \tilde{G} associated to $\tilde{\lambda}$ is defined for any $x, y \in E$, $0 \leq t \leq t^*(x)$, by

$$(15) \quad \tilde{G}(x, y, t) = \exp \left(- \int_0^t \tilde{\lambda}(x, y, s) ds \right).$$

Moreover, for any $0 \leq t < t^*(x)$, $\tilde{\lambda}(x, y, t) = -u'(t)/u(t)$ with

$$u(t) = \int_t^{t^*(x)} f\tilde{Q}(x, s, y) ds + G\tilde{Q}(x, t^*(x), y).$$

As a consequence, we have

$$\begin{aligned} \int_0^t \tilde{\lambda}(x, y, s) ds &= -\ln \left(\int_t^{t^*(x)} f\tilde{Q}(x, s, y) ds + G\tilde{Q}(x, t^*(x), y) \right) \\ &\quad + \ln \left(\int_0^{t^*(x)} f\tilde{Q}(x, s, y) ds + G\tilde{Q}(x, t^*(x), y) \right) \\ &= -\ln \left(\frac{\int_t^{t^*(x)} f\tilde{Q}(x, s, y) ds + G\tilde{Q}(x, t^*(x), y)}{\int_0^{t^*(x)} f\tilde{Q}(x, s, y) ds + G\tilde{Q}(x, t^*(x), y)} \right). \end{aligned}$$

Finally, together with (15),

$$(16) \quad \tilde{G}(x, y, t) = \frac{\int_t^{t^*(x)} f\tilde{Q}(x, s, y) ds + G\tilde{Q}(x, t^*(x), y)}{\int_0^{t^*(x)} f\tilde{Q}(x, s, y) ds + G\tilde{Q}(x, t^*(x), y)}.$$

In order to establish the expected result, we need to prove the equality

$$(17) \quad \eta_n(A \times B \times]t, +\infty[) = \int_{A \times B} \tilde{G}(x, y, t \wedge t^*(x)) \mathbf{1}_{\{0 \leq t < t^*(x)\}} \tilde{\nu}_n(dx \times dy),$$

for any $A, B \in \mathcal{B}(E)$ and $t \geq 0$. By (12), we have

$$\begin{aligned} \int_{A \times B} \tilde{G}(x, y, t \wedge t^*(x)) \mathbf{1}_{\{0 \leq t < t^*(x)\}} \tilde{\nu}_n(dx \times dy) \\ = \int_{A \times B} \tilde{G}(x, y, t \wedge t^*(x)) \mathbf{1}_{\{0 \leq t < t^*(x)\}} R(x, dy) \nu_n(dx). \end{aligned}$$

Thus, with (10) and (16), we obtain

$$\begin{aligned} \int_{A \times B} \tilde{G}(x, y, t \wedge t^*(x)) \mathbf{1}_{\{0 \leq t < t^*(x)\}} \tilde{\nu}_n(dx \times dy) \\ = \int_{A \times B} \mathbf{1}_{\{0 \leq t < t^*(x)\}} \left[\int_{t \wedge t^*(x)}^{t^*(x)} f \tilde{Q}(x, s, y) ds + G \tilde{Q}(x, t^*(x), y) \right] \mu(dy) \nu_n(dx) \\ = \int_A \mathbf{1}_{\{0 \leq t < t^*(x)\}} \left[\int_{t \wedge t^*(x)}^{t^*(x)} f \bar{Q}(x, s, B) ds + G \bar{Q}(x, t^*(x), B) \right] \nu_n(dx). \end{aligned}$$

Together with the expression (13) of η_n , this directly implies (17) and, therefore, the expected result. \square

Remark 3.9. Let n be an integer. On the strength of Jirina's theorem (see for instance Theorem 11.7 of [22]), there exists a kernel family $(\gamma_{x,y}(\cdot))_{(x,y) \in E^2}$ such that for any $A \times B \times \Gamma$ in $\mathcal{B}(E)^{\otimes 2} \otimes \mathcal{B}(\mathbf{R}_+)$, we have

$$\eta_n(A \times B \times \Gamma) = \int_{A \times B} \gamma_{x,y}(\Gamma) \tilde{\nu}_n(dx \times dy).$$

Let $x, y \in E^2$. By Proposition 3.8, $\gamma_{x,y}$ does not depend on n . Furthermore, we have the relation

$$\begin{aligned} \gamma_{x,y}([t, +\infty[) &= \mathbf{P}_{\nu_0}(S_1 > t \mid Z_0 = x, Z_1 = y) \\ &= \tilde{G}(x, y, t), \end{aligned}$$

if $t < t^*(x)$, by (17).

We also have the following continuous-time martingale property.

Corollary 3.10. Let i be an integer. The continuous-time process M^{i+1} given by,

$$(18) \quad \forall 0 \leq t < t^*(Z_i), \quad M^{i+1}(t) = N^{i+1}(t) - \int_0^t \tilde{\lambda}(Z_i, Z_{i+1}, u) \mathbf{1}_{\{S_{i+1} \geq u\}} du,$$

is a $(\sigma(Z_i, Z_{i+1}) \vee \mathcal{F}_t^{i+1})_{0 \leq t < t^*(Z_i)}$ -continuous-time martingale.

Proof. The proof is similar to the one of Lemma 2.2 of [6]. Indeed, the conditional jump rate of S_{i+1} given Z_i, Z_{i+1} is $\tilde{\lambda}(Z_i, Z_{i+1}, \cdot)$ by Proposition 3.8. \square

4. ERGODICITY

In this section, we focus our attention on the asymptotic behavior of the Markov chains $(Z_n)_{n \geq 0}$, $(Z_n, S_{n+1})_{n \geq 0}$ and $(Z_n, Z_{n+1}, S_{n+1})_{n \geq 0}$. Our main objective is to state that one can apply the ergodic theorem to these Markov chains. Before, we impose a new condition on the function t^* .

Assumption 4.1. $t^* : E \rightarrow \mathbf{R}_+$ is a bounded function.

Under this new assumption, we derive uniform lower bounds for the transition kernel R and $G\tilde{Q}(x, t^*(x), y)$.

Remark 4.2. Let $x, y \in E$ and $B \in \mathcal{B}(E)$. By (7) and (8), we have

$$(19) \quad G\tilde{Q}(x, t^*(x), y) \geq m_2,$$

where m_2 is given by

$$m_2 = m \exp \left(- \int_0^{\|t^*\|_\infty} M(s) ds \right) > 0.$$

In particular, m_2 do not depend on x and y . Together with (11), we have

$$(20) \quad R(x, B) \geq m_2 \mu(B).$$

According to the previous remark, one may state that the Markov chain $(Z_n)_{n \geq 0}$ is ergodic.

Proposition 4.3. We have the following statements:

- a) $(Z_n)_{n \geq 0}$ is μ -irreducible, aperiodic and admits a unique invariant measure, which we denote by ν .
- b) There exist $\rho > 1$ and $r > 0$ such that,

$$(21) \quad \forall n \geq 0, \sup_{\xi \in E} \|R^n(\xi, \cdot) - \nu\|_{TV} \leq r \rho^{-n},$$

where $\|\cdot\|_{TV}$ denotes the total variation norm (see [15] for the definition).

- c) The Markov chain $(Z_n)_{n \geq 0}$ is positive Harris-recurrent.

Proof. By definition and Remark 4.2, $(Z_n)_{n \geq 0}$ is μ -irreducible and aperiodic. In addition, the transition kernel R obviously satisfies Doeblin's condition (see [21, page 396] for instance),

$$\mu(B) > \epsilon \Rightarrow R(\xi, B) > m_2 \epsilon,$$

by (20). On the strength of Theorem 16.0.2 of [21], $(Z_n)_{n \geq 0}$ admits a unique invariant measure ν since it is aperiodic and (21) holds. In addition, from Theorem 4.3.3 of [15], $(Z_n)_{n \geq 0}$ is positive Harris-recurrent. \square

Now, we shall see that the sets with non-empty interior are charged by the invariant measure ν .

Remark 4.4. The transition kernel R admits a density according to the measure μ . As a consequence, the invariant measure ν and the auxiliary measure μ are equivalent in the light of Theorem 10.4.9 of [21]. This ensures that for any measurable set A with non-empty interior, $\nu(A) > 0$ by Assumptions 3.2.

The following lemma deals with the limits of the sequences $(\tilde{\nu}_n)_{n \geq 0}$ and $(\eta_n)_{n \geq 0}$.

Lemma 4.5. For any initial distribution $\nu_0 = \delta_{\{x\}}$, $x \in E$,

$$\lim_{n \rightarrow +\infty} \|\tilde{\nu}_n - \tilde{\nu}\|_{TV} = 0 \quad \text{and} \quad \lim_{n \rightarrow +\infty} \|\eta_n - \eta\|_{TV} = 0,$$

where the limit distributions $\tilde{\nu}$ and η are given by,

$$(22) \quad \forall A \times B \in \mathcal{B}(E)^{\otimes 2}, \tilde{\nu}(A \times B) = \int_A \nu(dx) R(x, B),$$

$$(23) \quad \forall A \times B \times \Gamma \in \mathcal{B}(E)^{\otimes 2} \otimes \mathcal{B}(\mathbf{R}_+), \eta(A \times B \times \Gamma) = \int_{A \times B \times \Gamma} \gamma_{x,y}(ds) \tilde{\nu}(dx \times dy).$$

Proof. Let g be a measurable function bounded by 1. By virtue of Fubini's theorem, we have

$$\left| \int_{E \times E} g(x, y) (\tilde{\nu}_n(dx \times dy) - \tilde{\nu}(dx \times dy)) \right| = \left| \int_E (\nu_n(dx) - \nu(dx)) \int_E g(x, y) R(x, dy) \right|,$$

from the expression of $\tilde{\nu}_n$ (12) and the definition of $\tilde{\nu}$ (22). Thus,

$$\left| \int_{E \times E} g(x, y) (\tilde{\nu}_n(dx \times dy) - \tilde{\nu}(dx \times dy)) \right| = \left| \int_E h(x) (\nu_n(dx) - \nu(dx)) \right|,$$

where $h : x \mapsto \int_E g(x, y) R(x, dy)$ is bounded by 1 because g is bounded by 1 and R is a transition kernel. Finally,

$$\|\tilde{\nu}_n - \tilde{\nu}\|_{TV} \leq \|\nu_n - \nu\|_{TV}.$$

One obtains the expected limit from (21). Now, we state the second limit. Let g' a measurable function bounded by 1. In the light of Remark 3.9,

$$\left| \int_{E \times E \times \mathbf{R}_+} g'(x, y, s) (\eta_n(dx \times dy \times ds) - \eta(dx \times dy \times ds)) \right| \leq \|h\|_{\infty} \|\tilde{\nu}_n - \tilde{\nu}\|_{TV},$$

by virtue of Fubini's theorem and with the function h given by,

$$\forall x, y \in E, h(x, y) = \int_{\mathbf{R}_+} g'(x, y, s) \gamma_{x,y}(ds).$$

As h is bounded by 1, we have

$$\left| \int_{E \times E \times \mathbf{R}_+} g'(x, y, s) (\eta_n(dx \times dy \times ds) - \eta(dx \times dy \times ds)) \right| \leq \|\tilde{\nu}_n - \tilde{\nu}\|_{TV}.$$

We previously established that $\|\tilde{\nu}_n - \tilde{\nu}\|_{TV}$ tends to 0, thus $\|\eta_n - \eta\|_{TV}$ tends to 0 too. \square

The previous lemma induces the following result.

Proposition 4.6. *We have the following statements:*

- a) $(Z_n, Z_{n+1})_{n \geq 0}$ (resp. $(Z_n, Z_{n+1}, S_{n+1})_{n \geq 0}$) is $\tilde{\nu}$ -irreducible (resp. η -irreducible).
- b) $(Z_n, Z_{n+1})_{n \geq 0}$ and $(Z_n, Z_{n+1}, S_{n+1})_{n \geq 0}$ are positive Harris-recurrent and aperiodic Markov chains.
- c) $\tilde{\nu}$ (resp. η) is the unique invariant measure of the chain $(Z_n, Z_{n+1})_{n \geq 0}$ (resp. of the chain $(Z_n, Z_{n+1}, S_{n+1})_{n \geq 0}$).

Proof. This result is a consequence of Lemma 4.5. The proof is similar to the one of Proposition 4.2 given in [6]. \square

According to the previous discussion, the Markov chains $(Z_n)_{n \geq 0}$, $(Z_n, Z_{n+1})_{n \geq 0}$ and $(Z_n, Z_{n+1}, S_{n+1})_{n \geq 0}$ are positive Harris-recurrent. As a consequence, one may apply the ergodic theorem to these Markov chains (Theorem 17.1.7 of [21]).

5. ESTIMATION OF \tilde{l}

In this section, we focus our attention on the estimation of the function \tilde{l} , which is an approximation of the jump rate $\tilde{\lambda}$. The precise definition of \tilde{l} may be found in Lemma 5.7. The estimation of \tilde{l} is a keystone in our procedure for estimating the conditional density f . The main result of this section states in Proposition 5.13. First, we shall impose some additional conditions on t^* , f and \tilde{Q} .

Assumptions 5.1.

- a) f is bounded.
- b) t^* is continuous.
- c) \tilde{Q} is bounded.

Two technical lemmas about lower and upper bounds are now presented. These results will be useful in all the sequel.

Lemma 5.2. *Let A and B two measurable sets, with A relatively compact and such that $\overline{A} \cap \partial E = \emptyset$. Thus,*

$$\inf_{\xi \in A} t^*(\xi) > 0.$$

In this case, we denote $t^(A) = \inf_{\xi \in A} t^*(\xi)$. In addition, for any $0 \leq t < t^*(A)$,*

$$(24) \quad \inf_{x \in A, y \in B} \tilde{G}(x, y, t) > 0.$$

Proof. For the first inequality, one may refer to the proof of Lemma 4.7 given in [6]. For the second inequality, we have from (16) and (19), for any $x, y \in E$, $0 \leq t < t^*(x)$,

$$\begin{aligned} \tilde{G}(x, y, t) &\geq \frac{G\tilde{Q}(x, t^*(x), y)}{\int_0^{t^*(x)} f\tilde{Q}(x, s, y)ds + G\tilde{Q}(x, t^*(x), y)} \\ &\geq \frac{m_2}{(\|t^*\|_\infty \|f\|_\infty + 1)\|\tilde{Q}\|_\infty}, \end{aligned}$$

because f , \tilde{Q} and t^* are bounded (according to Assumptions 5.1 and Assumption 4.1). This achieves the proof. \square

One may also prove that the jump rate $\tilde{\lambda}$ is a bounded function.

Lemma 5.3. *Let $x, y \in E$ and $0 \leq t \leq t^*(x)$. Thus,*

$$|\tilde{\lambda}(x, y, t)| \leq \frac{\|f\|_\infty \|\tilde{Q}\|_\infty}{m_2}.$$

Proof. By (14), we have

$$\begin{aligned} \tilde{\lambda}(x, y, t) &= \frac{f\tilde{Q}(x, t, y)}{\int_t^{t^*(x)} f\tilde{Q}(x, s, y)ds + G\tilde{Q}(x, t^*(x), y)} \\ &\leq \frac{f\tilde{Q}(x, t, y)}{G\tilde{Q}(x, t^*(x), y)}. \end{aligned}$$

Therefore, with (19) and since f and \tilde{Q} are bounded, one immediately obtains the expected result. \square

Let us denote by \mathcal{C}_1 the following set,

$$\mathcal{C}_1 = \{B \in \mathcal{B}(E) : \overset{\circ}{B} \neq \emptyset\}.$$

In addition, \mathcal{C}_2 is defined by

$$\mathcal{C}_2 = \{A \in \mathcal{C}_1 : A \text{ relatively compact set such that } \overline{A} \cap \partial E = \emptyset\}.$$

Remark 5.4. Let $A, B \in \mathcal{C}_1$. Then, $\tilde{\nu}(A \times B) > 0$. Indeed,

$$\tilde{\nu}(A \times B) = \int_A R(x, B) \nu(dx).$$

For any x , $R(x, B) \geq m_2 \mu(B)$ by (20), thus,

$$\tilde{\nu}(A \times B) \geq m_2 \mu(B) \nu(A).$$

One may conclude because $\nu(A) > 0$ by Remark 4.4 and $\mu(B) > 0$ since B is a set with non-empty interior (see Assumptions 3.2).

Let $A, B \in \mathcal{C}_1$. For any integer n , one defines the process $Y_n(A, B, \cdot)$ by,

$$(25) \quad \forall 0 \leq t < t^*(A), Y_n(A, B, t) = \sum_{i=0}^{n-1} \mathbf{1}_{\{S_{i+1} \geq t\}} \mathbf{1}_{\{Z_i \in A\}} \mathbf{1}_{\{Z_{i+1} \in B\}}.$$

The following results deal with the asymptotic properties of Y_n .

Lemma 5.5. Let $A \in \mathcal{C}_2$ and $B \in \mathcal{C}_1$. Then, for any $x \in E$,

$$\forall 0 \leq t < t^*(A), \frac{Y_n(A, B, t)}{n} \rightarrow \int_{A \times B} \tilde{G}(x, y, t) \tilde{\nu}(dx \times dy) \mathbf{P}_x\text{-a.s.},$$

when n goes to infinity. In addition, this limit is strictly positive.

Proof. By virtue of the ergodic theorem applied to the chain $(Z_n, Z_{n+1}, S_{n+1})_{n \geq 0}$,

$$\frac{1}{n} Y_n(A, B, t) \rightarrow \eta(A \times B \times [t, +\infty[) \mathbf{P}_x\text{-a.s.}$$

Together with (23) and Remark 3.9, we have for any $0 \leq t < t^*(A)$,

$$\frac{Y_n(A, B, t)}{n} \rightarrow \int_{A \times B} \tilde{G}(x, y, t) \tilde{\nu}(dx \times dy) \mathbf{P}_x\text{-a.s.},$$

when n goes to infinity. Since $\inf_{x \in A, y \in B} \tilde{G}(x, y, t) > 0$ by (24) and $\tilde{\nu}(A \times B) > 0$ by Remark 5.4, the limit is strictly positive. \square

For any $A \in \mathcal{C}_2$, $B \in \mathcal{C}_1$ and $0 \leq t < t^*(A)$, we define the process $Y_n(A, B, t)^+$, which is the generalized inverse of $Y_n(A, B, t)$, by

$$(26) \quad Y_n(A, B, t)^+ = \begin{cases} 0 & \text{if } Y_n(A, B, t) = 0, \\ \frac{1}{Y_n(A, B, t)} & \text{else.} \end{cases}$$

Lemma 5.6. *Let $A \in \mathcal{C}_2$, $B \in \mathcal{C}_1$, $0 \leq t < t^*(A)$ and $x \in E$. Then, for any integer n ,*

$$Y_n(A, B, t)^+ \leq 1 \quad \mathbf{P}_x\text{-a.s.}$$

and, as n goes to infinity,

$$\begin{aligned} Y_n(A, B, t)^+ &\longrightarrow 0 \quad \mathbf{P}_x\text{-a.s.}, \\ \mathbf{1}_{\{Y_n(A, B, t)=0\}} &\longrightarrow 0 \quad \mathbf{P}_x\text{-a.s.}, \\ \int_0^t \mathbf{1}_{\{Y_n(A, B, s)=0\}} ds &\longrightarrow 0 \quad \mathbf{P}_x\text{-a.s.} \end{aligned}$$

Proof. This result is a corollary of Lemma 5.5. One may find a similar proof in [6], Lemma 4.11. \square

The function \tilde{l} is defined in the following proposition.

Proposition 5.7. *Let $A \in \mathcal{C}_2$ and $B \in \mathcal{C}_1$. Let $0 \leq t < t^*(A)$ and $x \in E$. When n goes to infinity,*

$$Y_n(A, B, t)^+ \sum_{i=0}^{n-1} \tilde{\lambda}(Z_i, Z_{i+1}, t) \mathbf{1}_{\{Z_i \in A\}} \mathbf{1}_{\{Z_{i+1} \in B\}} \mathbf{1}_{\{S_{i+1} \geq t\}} \longrightarrow \tilde{l}(A, B, t) \quad \mathbf{P}_x\text{-a.s.},$$

where

$$(27) \quad \tilde{l}(A, B, t) = \frac{\int_{A \times B} \tilde{\lambda}(u, v, t) \tilde{G}(u, v, t) \tilde{\nu}(du \times dv)}{\int_{A \times B} \tilde{G}(u, v, t) \tilde{\nu}(du \times dv)}.$$

In addition, $\tilde{l}(A, B, \cdot)$ is continuous on $[0, t^*(A)[$.

Proof. This is an application of the ergodic theorem to the chain $(Z_n, Z_{n+1}, S_{n+1})_{n \geq 0}$. One may refer the reader to the proof of Proposition 4.12 of [6], which is similar. $\tilde{l}(A, B, \cdot)$ is continuous because $\tilde{\lambda}$ and \tilde{G} are continuous and bounded. Therefore, one may apply the theorem of continuity under the integral sign. \square

Proposition 5.8. *Let n be an integer and $1 \leq i \leq n$. For each $j \neq i$, let $t_j \geq 0$ and t be a positive real number. Then, we have*

$$\bigvee_{j \neq i} \mathcal{F}_{t_j}^j \perp_{\mathcal{G}_n} \mathcal{F}_t^i \quad \text{and} \quad \mathcal{F}_t^i \perp_{\sigma(Z_{i-1}, Z_i)} \mathcal{G}_n.$$

We deduce from Proposition 6.8 of [17] this direct corollary. For any $s < t$, $0 \leq i \leq n-1$,

$$\bigvee_{j \neq i+1} \mathcal{F}_s^j \perp_{\mathcal{G}_n \vee \mathcal{F}_s^{i+1}} \mathcal{F}_t^{i+1} \quad \text{and} \quad \mathcal{F}_t^{i+1} \perp_{\sigma(Z_i, Z_{i+1}) \vee \mathcal{F}_s^{i+1}} \mathcal{G}_n.$$

Proof. The technical proof is deferred in Appendix A. \square

Theorem 5.9. *Let $A \in \mathcal{C}_2$, $B \in \mathcal{C}_1$ and n be an integer. The process $M_n(A, B, \cdot)$, defined for any $0 \leq t < t^*(A)$, by*

$$M_n(A, B, t) = \sum_{i=0}^{n-1} M^{i+1}(t) \mathbf{1}_{\{Z_i \in A\}} \mathbf{1}_{\{Z_{i+1} \in B\}},$$

is a $(\mathcal{G}_n \vee \bigvee_{i=0}^{n-1} \mathcal{F}_t^{i+1})_{0 \leq t < t^(A)}$ -continuous-time martingale.*

Proof. One may recall that the process M^{i+1} has already been defined by (18). The result is a corollary of Proposition 5.8 and Corollary 3.10. The arguments are given in the proof of Theorem 4.13 of [6] \square

Lemma 5.10. *Let $A \in \mathcal{C}_2$, $B \in \mathcal{C}_1$ and n be an integer. The process $\widetilde{M}_n(A, B, \cdot)$, given for any $0 \leq t < t^*(A)$, by*

$$\widetilde{M}_n(A, B, t) = \int_0^t Y_n(A, B, s)^+ dM_n(A, B, s),$$

is a continuous-time martingale, whose predictable variation process $\langle \widetilde{M}_n(A, B) \rangle$ satisfies, for any $x \in E$,

$$\forall 0 \leq t < t^*(A), \langle \widetilde{M}_n(A, B) \rangle(t) \rightarrow 0 \text{ } \mathbf{P}_x\text{-a.s. when } n \rightarrow +\infty.$$

Proof. This is a consequence of Lemma 5.3, Lemma 5.6 and Theorem 5.9. A similar proof may be found in [6], Lemma 4.14. \square

Remark 5.11. *This immediately induces that for any $A \in \mathcal{C}_2$, $B \in \mathcal{C}_1$ and $0 \leq t < t^*(A)$,*

$$\sup_{0 \leq s \leq t} |\widetilde{M}_n(A, B, s)| \xrightarrow{\mathbf{P}_x} 0,$$

for any $x \in E$, by virtue of Lenglart's inequality. A reference may be found in the book [5], II.5.2.1. Lenglart's inequality.

In the sequel, we are interested in the estimation of the function $\widetilde{l}(A, B, \cdot)$, defined by (27), for any $A \in \mathcal{C}_2$ and $B \in \mathcal{C}_1$. First, we shall estimate the function $\widetilde{L}(A, B, \cdot)$ defined by,

$$\forall 0 \leq t < t^*(A), \widetilde{L}(A, B, t) = \int_0^t \widetilde{l}(A, B, s) ds.$$

We consider the Nelson-Aalen type estimator $\widehat{\widetilde{L}}_n(A, B, \cdot)$ given by,

$$\forall 0 \leq t < t^*(A), \widehat{\widetilde{L}}_n(A, B, t) = \int_0^t Y_n(A, B, s)^+ dN_n(A, B, s),$$

where $Y_n(A, B, s)^+$ has already been defined by (26) and where $N_n(A, B, \cdot)$ is the counting process given by,

$$\forall 0 \leq t < t^*(A), N_n(A, B, t) = \sum_{i=0}^{n-1} \mathbf{1}_{\{S_{i+1} \leq t\}} \mathbf{1}_{\{Z_i \in A\}} \mathbf{1}_{\{Z_{i+1} \in B\}}.$$

According to foregoing, the present framework is exactly the same one as in [6].

Proposition 5.12. *Let $A \in \mathcal{C}_2$, $B \in \mathcal{C}_1$, $0 < t < t^*(A)$ and $x \in E$. Then,*

$$\sup_{0 \leq s \leq t} \left| \widehat{\widetilde{L}}_n(A, B, s) - \widetilde{L}(A, B, s) \right| \xrightarrow{\mathbf{P}_x} 0 \text{ when } n \rightarrow +\infty.$$

Proof. This is a corollary of Lemma 5.10 and Proposition 5.7. One may refer the interested reader to the proofs of Proposition 4.16 and Theorem 4.17 of [6], which use similar arguments. \square

We focus on smoothing this estimator in order to provide an estimate of $\tilde{l}(A, B, \cdot)$. Let K be a continuous kernel whose support is $[-1, 1]$, $b > 0$ and $0 < t < t^*(A)$. We denote,

$$\forall 0 \leq s \leq t, \hat{l}_{n,b,t}(A, B, s) = \frac{1}{b} \int_0^t K\left(\frac{s-u}{b}\right) d\hat{L}_n(A, B, u).$$

The following result deals with the asymptotic behavior of this estimator.

Proposition 5.13. *Let $A \in \mathcal{C}_2$ and $B \in \mathcal{C}_1$. Let $0 < r_1 < r_2 < t < t^*(A)$ and $x \in E$. There exists a sequence $(\beta_n)_{n \geq 0}$ (depending on A, B and t) which almost surely tends to 0, such that*

$$\sup_{r_1 \leq s \leq r_2} \left| \hat{l}_{n,\beta_n,t}(A, B, s) - \tilde{l}(A, B, s) \right| \xrightarrow{\mathbf{P}_x} 0 \text{ when } n \rightarrow +\infty.$$

Proof. The proof relies on the arguments provided in the one of Proposition 4.23 in [6]. \square

As a conclusion of this section, we provided a consistent estimator of the function \tilde{l} . This is prominent in our procedure for estimating f .

6. ESTIMATION OF THE CONDITIONAL DENSITY f

In this section, we shall present a consistent estimator of the conditional density of interest f . Our method consists in studying the link between \tilde{l} and f . We will use this relation to provide an estimate of f , and prove a result of uniform convergence in Theorem 6.11. One considers the distance d_2 on \mathcal{E}^2 induced by the distance d and the Euclidean norm on \mathbf{R}^2 . For any $(x, y), (u, v) \in \mathcal{E}^2$,

$$d_2((x, y), (u, v)) = \sqrt{d(x, u)^2 + d(y, v)^2}.$$

Let us first impose regularity conditions on f , $f\tilde{Q}$ and $G\tilde{Q}$ from which we shall deduce that $\tilde{\lambda}$ is Lipschitz (see Lemma 6.3). It will allow us to approximate $\tilde{\lambda}$ by \tilde{l} (see Lemma 6.6).

Assumptions 6.1.

a) *There exists a constant $[f\tilde{Q}]_{Lip} > 0$ such that, for any $x, y, u, v \in E$ and for any $0 \leq t < t^*(x) \wedge t^*(u)$,*

$$(28) \quad \left| f\tilde{Q}(x, t, y) - f\tilde{Q}(u, t, v) \right| \leq [f\tilde{Q}]_{Lip} d_2((x, y), (u, v)).$$

b) *There exists a constant $[f]_{Lip} > 0$ such that,*

$$\forall x, y \in E, \forall t \geq 0, |f(x, t) - f(y, t)| \leq [f]_{Lip} d(x, y).$$

c) *For any $\xi \in E$, $t \geq 0$, $f(\xi, t) > 0$.*

d) *There exists a constant $[G\tilde{Q}]_{Lip} > 0$ such that, for any $x, y, u, v \in E$,*

$$(29) \quad \left| G\tilde{Q}(x, t^*(x), y) - G\tilde{Q}(u, t^*(u), v) \right| \leq [G\tilde{Q}]_{Lip} d_2((x, y), (u, v)).$$

For the sake of clarity, we denote for any $x, y \in E$ and $0 \leq t \leq t^*(x)$,

$$(30) \quad H(x, y, t) = \int_t^{t^*(x)} f\tilde{Q}(x, s, y) ds + G\tilde{Q}(x, t^*(x), y).$$

Therefore, by (14), we have

$$H(x, y, t)\tilde{\lambda}(x, y, t) = f\tilde{Q}(x, t, y).$$

We shall establish some properties of H .

Lemma 6.2. *We have the following statements:*

- a) $\|H\|_\infty < +\infty$.
- b) For any $x, y \in E$, $0 \leq t \leq t^*(x)$, $H(x, y, t) \geq m_2$.
- c) Furthermore, there exists a constant $[H]_{Lip} > 0$ such that, for any $x, y, u, v \in E$ and $0 \leq t \leq t^*(x) \wedge t^*(u)$,

$$(31) \quad |H(x, y, t) - H(u, v, t)| \leq [H]_{Lip} d_2((x, y), (u, v)).$$

Proof. First, by (30), we have

$$H(x, y, t) \leq \int_0^{t^*(x)} f\tilde{Q}(x, s, y) ds + G\tilde{Q}(x, t^*(x), y).$$

G is bounded by 1. In addition, f , t^* and \tilde{Q} are bounded functions according to Assumptions 5.1, Assumption 4.1 and Assumptions 6.1. Thus,

$$H(x, y, t) \leq \|\tilde{Q}\|_\infty (\|t^*\|_\infty \|f\|_\infty + 1).$$

Now, we prove that m_2 is a lower bound of H . Indeed, by (30) again, for any $x, y \in E$ and $0 \leq t \leq t^*(x)$,

$$(32) \quad H(x, y, t) \geq G\tilde{Q}(x, t^*(x), y) \geq m_2.$$

Finally, we state that H is Lipschitz. By (30) again and by the triangle inequality,

$$\begin{aligned} & |H(x, y, t) - H(u, v, t)| \\ & \leq \int_t^{t^*(x) \vee t^*(u)} |f\tilde{Q}(u, s, v) - f\tilde{Q}(x, s, y)| ds + |G\tilde{Q}(u, t^*(u), v) - G\tilde{Q}(x, t^*(x), y)|. \end{aligned}$$

Thus, since the product $f\tilde{Q}$ is Lipschitz by (28), we have

$$\begin{aligned} & |H(x, y, t) - H(u, v, t)| \\ & \leq \|t^*\|_\infty [f\tilde{Q}]_{Lip} d_2((x, y), (u, v)) + |G\tilde{Q}(u, t^*(u), v) - G\tilde{Q}(x, t^*(x), y)|. \end{aligned}$$

Together with (29), we obtain (31) with

$$[H]_{Lip} = \|t^*\|_\infty [f\tilde{Q}]_{Lip} + [G\tilde{Q}]_{Lip},$$

showing the three statements. □

Now, one may state that $\tilde{\lambda}$ is also Lipschitz.

Lemma 6.3. *There exists a constant $[\tilde{\lambda}]_{Lip} > 0$ such that, for any $x, y, u, v \in E$ and $0 \leq t < t^*(x) \wedge t^*(u)$,*

$$|\tilde{\lambda}(x, y, t) - \tilde{\lambda}(u, v, t)| \leq [\tilde{\lambda}]_{Lip} d_2((x, y), (u, v)).$$

Proof. From (14) and (30), we have

$$\left| \tilde{\lambda}(x, y, t) - \tilde{\lambda}(u, v, t) \right| = \frac{|f\tilde{Q}(x, t, y)H(u, v, t) - f\tilde{Q}(u, t, v)H(x, y, t)|}{H(u, v, t)H(x, y, t)}.$$

Therefore, together with (32), we obtain

$$\left| \tilde{\lambda}(x, y, t) - \tilde{\lambda}(u, v, t) \right| \leq \frac{1}{m_2^2} |f\tilde{Q}(x, t, y)H(u, v, t) - f\tilde{Q}(u, t, v)H(x, y, t)|.$$

By the triangle inequality, this induces that

$$\begin{aligned} & \left| \tilde{\lambda}(x, y, t) - \tilde{\lambda}(u, v, t) \right| \\ & \leq \frac{1}{m_2^2} \left(|f\tilde{Q}(x, t, y)H(u, v, t) - f\tilde{Q}(x, t, y)H(x, y, t)| \right. \\ & \quad \left. + |f\tilde{Q}(x, t, y)H(x, y, t) - f\tilde{Q}(u, t, v)H(x, y, t)| \right) \\ & \leq \frac{1}{m_2^2} \left(\|f\|_\infty \|\tilde{Q}\|_\infty |H(u, v, t) - H(x, y, t)| + \|H\|_\infty |f\tilde{Q}(x, t, y) - f\tilde{Q}(u, t, v)| \right). \end{aligned}$$

$f\tilde{Q}$ is Lipschitz in the light of Assumptions 6.1 and H is Lipschitz by Lemma 6.2. Hence,

$$\left| \tilde{\lambda}(x, y, t) - \tilde{\lambda}(u, v, t) \right| \leq [\tilde{\lambda}]_{Lip} d_2((x, y), (u, v)),$$

where $[\tilde{\lambda}]_{Lip}$ is given by

$$[\tilde{\lambda}]_{Lip} = \frac{\|f\|_\infty \|\tilde{Q}\|_\infty [H]_{Lip} + \|H\|_\infty [f\tilde{Q}]_{Lip}}{m_2^2}.$$

This achieves the proof. \square

We impose a new condition on \overline{Q} .

Assumption 6.4. *There exist $p \geq 2$ and a family $\{B_1, \dots, B_p\}$, $B_k \in \mathcal{C}_1$ for any k , such that*

$$\max_{1 \leq k \leq p} \text{diam } B_k < +\infty \quad \text{and} \quad \forall x \in E, \forall t \geq 0, \sum_{k=1}^p \overline{Q}(x, t, B_k) = 1.$$

Without loss of generality, B_1 is assumed to be the set with the largest diameter.

If E is bounded, then Assumption 6.4 is obviously satisfied.

Remark 6.5. *Under this assumption, one may state a new property of the measure μ . By (2), we have*

$$\overline{Q}(x, t, B_k) = \int_{B_k} \tilde{Q}(x, t, y) \mu(dy) \geq m \mu(B_k).$$

Summing on k yields to

$$(33) \quad \sum_{k=1}^p \mu(B_k) \leq \frac{1}{m},$$

by virtue of Assumption 6.4.

In the following lemma, we establish that \tilde{l} is an approximation of $\tilde{\lambda}$.

Lemma 6.6. *Let $A \in \mathcal{C}_2$, $1 \leq k \leq p$, $x \in A$ and $y \in B_k$. Let $0 \leq t < t^*(A)$. Then,*

$$(34) \quad |\tilde{\lambda}(x, y, t) - \tilde{l}(A, B_k, t)| \leq [\tilde{\lambda}]_{Lip} \text{diam}_2 A \times B_1.$$

Proof. By (27), we have

$$\begin{aligned} |\tilde{\lambda}(x, y, t) - \tilde{l}(A, B_k, t)| &\leq \left| \tilde{\lambda}(x, y, t) - \frac{\int_{A \times B_k} \tilde{\lambda}(u, v, t) \tilde{G}(u, v, t) \tilde{\nu}(du \times dv)}{\int_{A \times B_k} \tilde{G}(u, v, t) \tilde{\nu}(du \times dv)} \right| \\ &\leq \frac{\int_{A \times B_k} |\tilde{\lambda}(x, y, t) - \tilde{\lambda}(u, v, t)| \tilde{G}(u, v, t) \tilde{\nu}(du \times dv)}{\int_{A \times B_k} \tilde{G}(u, v, t) \tilde{\nu}(du \times dv)} \\ &\leq [\tilde{\lambda}]_{Lip} \text{diam}_2 A \times B_k, \end{aligned}$$

by virtue of Lemma 6.3. This yields to the expected result, since B_1 is assumed to be the set with the largest diameter. \square

An approximation of the density f is presented in the following result.

Proposition 6.7. *Let $A \in \mathcal{C}_2$ and $\xi \in A$. For any $0 \leq t < t^*(A)$,*

$$\begin{aligned} \left| f(\xi, t) - \frac{1}{\nu(A)} \sum_{k=1}^p \tilde{l}(A, B_k, t) \int_{A \times B_k} H(x, y, t) \mu(dy) \nu(dx) \right| \\ \leq \frac{1}{m} [\tilde{\lambda}]_{Lip} \|H\|_{\infty} \text{diam}_2 A \times B_1 + [f]_{Lip} \text{diam } A. \end{aligned}$$

Proof. First, we show that

$$(35) \quad \left| \frac{1}{\nu(A)} \int_A f(x, t) \nu(dx) - \sum_{k=1}^p \tilde{l}(A, B_k, t) \int_{A \times B_k} H(x, y, t) \mu(dy) \nu(dx) \right| \leq \frac{1}{m} [\tilde{\lambda}]_{Lip} \|H\|_{\infty} \text{diam}_2 A \times B_1.$$

Let $1 \leq k \leq p$, $x \in A$ and $y \in B_k$. Multiplying (34) by $H(x, y, t)$, we obtain

$$|f\tilde{Q}(x, t, y) - \tilde{l}(A, B_k, t)H(x, y, t)| \leq [\tilde{\lambda}]_{Lip} \|H\|_{\infty} \text{diam}_2 A \times B_1,$$

because H is bounded on the strength of Lemma 6.2. We integrate on B_k according to $\mu(dy)$. We obtain

$$\left| f\tilde{Q}(x, t, B_k) - \tilde{l}(A, B_k, t) \int_{B_k} H(x, y, t) \mu(dy) \right| \leq \mu(B_k) [\tilde{\lambda}]_{Lip} \|H\|_{\infty} \text{diam}_2 A \times B_1.$$

Summing on k between 1 and p yields to

$$\left| f(x, t) - \sum_{k=1}^p \tilde{l}(A, B_k, t) \int_{B_k} H(x, y, t) \mu(dy) \right| \leq \frac{1}{m} [\tilde{\lambda}]_{Lip} \|H\|_{\infty} \text{diam}_2 A \times B_1,$$

with (33). Finally, we integrate on A according to $\nu(dx)$.

$$(36) \quad \left| \int_A f(x, t) \nu(dx) - \sum_{k=1}^p \tilde{l}(A, B_k, t) \int_{A \times B_k} H(x, y, t) \nu(dx) \mu(dy) \right| \leq \frac{\nu(A)}{m} [\tilde{\lambda}]_{Lip} \|H\|_{\infty} \text{diam}_2 A \times B_1.$$

Doing the ratio of each term of (36) by $\nu(A)$ leads to (35). On the other hand, we have

$$(37) \quad \left| f(\xi, t) - \frac{\int_A f(x, t) \nu(dx)}{\nu(A)} \right| \leq [f]_{Lip} \text{diam } A,$$

because f is Lipschitz by Assumptions 6.1. Finally, (35) and (37) give the proof by virtue of the triangle inequality. \square

Let $A \in \mathcal{C}_2$, $\xi \in A$ and $0 \leq t < t^*(A)$. For estimating $f(\xi, t)$, we need to estimate $\tilde{l}(A, B_k, t)$ and

$$\frac{1}{\nu(A)} \int_{A \times B_k} H(x, y, t) \nu(dx) \mu(dy),$$

for each $1 \leq k \leq p$. We shall see that this quantity may be seen as a conditional probability.

Proposition 6.8. *Let $A \in \mathcal{C}_2$, $1 \leq k \leq p$ and $0 \leq t < t^*(A)$. Then,*

$$\int_{A \times B_k} H(x, y, t) \nu(dx) \mu(dy) = \mathbf{P}_\nu(S_1 > t, Z_1 \in B_k, Z_0 \in A).$$

This immediately induces that

$$\frac{1}{\nu(A)} \int_{A \times B_k} H(x, y, t) \nu(dx) \mu(dy) = \mathbf{P}_\nu(S_1 > t, Z_1 \in B_k | Z_0 \in A).$$

One may recall that $\nu(A)$ is a strictly positive number according to Remark 4.4.

Proof. By (23) and Remark 3.9, we have for any $t < t^*(A)$,

$$(38) \quad \begin{aligned} \mathbf{P}_\nu(S_1 > t, Z_1 \in B_k, Z_0 \in A) &= \eta(A \times B_k \times [t, +\infty[) \\ &= \int_{A \times B_k} \tilde{G}(x, y, t) \tilde{\nu}(dx \times dy). \end{aligned}$$

Hence,

$$\mathbf{P}_\nu(S_1 > t, Z_1 \in B_k, Z_0 \in A) = \int_{A \times B_k} \tilde{G}(x, y, t) \nu(dx) R(x, dy),$$

by (22). Thus, by the expression of R (9), the definition of \tilde{G} (16) and the definition of H (30), we have

$$\tilde{G}(x, y, t) R(x, dy) = H(x, y, t) \mu(dy).$$

Therefore,

$$(39) \quad \mathbf{P}_\nu(S_1 > t, Z_1 \in B_k, Z_0 \in A) = \int_{A \times B_k} H(x, y, t) \nu(dx) \mu(dy),$$

showing the result. \square

Let $A \in \mathcal{C}_2$, $1 \leq k \leq p$ and $0 \leq t < t^*(A)$. One may estimate the conditional probability

$$\frac{1}{\nu(A)} \int_{A \times B_k} H(x, y, t) \nu(dx) \mu(dy) = \mathbf{P}_\nu(S_1 > t, Z_1 \in B_k | Z_0 \in A),$$

by its empirical version,

$$(40) \quad \hat{p}_n(A, B_k, t) = \frac{\sum_{i=0}^{n-1} \mathbf{1}_{\{S_{i+1} > t\}} \mathbf{1}_{\{Z_{i+1} \in B_k\}} \mathbf{1}_{\{Z_i \in A\}}}{\sum_{i=0}^{n-1} \mathbf{1}_{\{Z_i \in A\}}}.$$

We shall prove the uniform convergence of this estimator in Proposition 6.10. First, we establish the pointwise convergence.

Lemma 6.9. *Let $A \in \mathcal{C}_2$ and $1 \leq k \leq p$. For any $0 \leq t < t^*(A)$, $x \in E$, we have as n goes to infinity,*

$$\left| \widehat{p}_n(A, B_k, t) - \frac{1}{\nu(A)} \int_{A \times B_k} H(\xi, y, t) \nu(d\xi) \mu(dy) \right| \rightarrow 0 \quad \mathbf{P}_x\text{-a.s.}$$

Proof. One applies the ergodic theorem to $(Z_n, Z_{n+1}, S_{n+1})_{n \geq 0}$. Thus, when n goes to infinity,

$$\frac{1}{n} \sum_{i=0}^{n-1} \mathbf{1}_{\{Z_i \in A\}} \mathbf{1}_{\{Z_{i+1} \in B_k\}} \mathbf{1}_{\{S_{i+1} > t\}} \rightarrow \eta(A \times B_k \times]t, +\infty[) \quad \mathbf{P}_x\text{-a.s.}$$

Together with (38) and (39),

$$(41) \quad \frac{1}{n} \sum_{i=0}^{n-1} \mathbf{1}_{\{Z_i \in A\}} \mathbf{1}_{\{Z_{i+1} \in B_k\}} \mathbf{1}_{\{S_{i+1} > t\}} \rightarrow \int_{A \times B_k} H(x, y, t) \nu(dx) \mu(dy) \quad \mathbf{P}_x\text{-a.s.}$$

In addition, by applying the ergodic theorem to the Markov chain $(Z_n)_{n \geq 0}$, we have

$$(42) \quad \frac{1}{n} \sum_{i=0}^{n-1} \mathbf{1}_{\{Z_i \in A\}} \rightarrow \nu(A) \quad \mathbf{P}_x\text{-a.s.}$$

Combining (40), (41) and (42), we show the expected result. \square

Now, we state that $\widehat{p}_n(A, B_k, \cdot)$ is a consistent estimator.

Proposition 6.10. *Let $A \in \mathcal{C}_2$, $1 \leq k \leq p$ and $0 < t < t^*(A)$. For any $x \in E$, we have when n goes to infinity,*

$$\sup_{0 \leq s \leq t} \left| \widehat{p}_n(A, B_k, s) - \frac{1}{\nu(A)} \int_{A \times B_k} H(\xi, y, s) \nu(d\xi) \mu(dy) \right| \rightarrow 0 \quad \mathbf{P}_x\text{-a.s.}$$

Proof. One considers the function β defined by,

$$\forall 0 \leq t < t^*(A), \quad \beta(t) = \frac{1}{\nu(A)} \int_{A \times B_k} H(\xi, y, t) \nu(d\xi) \mu(dy).$$

β implicitly depends on A and k . From Lemma 6.9, for any s , $\mathbf{P}_x(\Upsilon_s) = 1$ where the set Υ_s is defined by

$$\Upsilon_s = \{\widehat{p}_n(A, B_k, s) \rightarrow \beta(s)\}.$$

First, we prove that β is a strictly decreasing function. By (30) and by virtue of the theorem of derivation under the integral sign, β' satisfies,

$$\begin{aligned} \forall 0 \leq s \leq t, \quad \beta'(s) &= -\frac{1}{\nu(A)} \int_A f(\xi, s) \overline{Q}(\xi, s, B_k) \nu(d\xi) \\ &\leq -\frac{\mu(B_k)m}{\nu(A)} \int_A f(\xi, s) \nu(d\xi), \end{aligned}$$

by (2). According to Assumptions 6.1, f is strictly positive. Thus, $\int_A f(\xi, s) \nu(d\xi) > 0$ because $\nu(A) > 0$. As a consequence, $\beta'(s) < 0$ for any s , and β is, therefore, strictly

decreasing. In particular, this is a one-to-one correspondence mapping from $[0, t]$ into $[a, b]$, where

$$[a, b] = [\beta(t), \mathbf{P}_\nu(Z_1 \in B_k | Z_0 \in A)],$$

since, in the light of Proposition 6.8,

$$\beta(0) = \mathbf{P}_\nu(Z_1 \in B_k | Z_0 \in A).$$

For any couple (l, m) of integers, with $0 \leq l \leq m$, let us consider

$$X(l, m) = \beta^{-1} \left(a + \frac{(m-l)(b-a)}{m} \right).$$

By construction, we have for any $0 \leq l \leq m-1$,

$$\beta(X(l+1, m)) - \beta(X(l, m)) = \frac{b-a}{m}.$$

Hence, for $s \in [X(l, m), X(l+1, m)[$ with $0 \leq l \leq m-2$, or for $s \in [X(m-1, m), t]$, we have

$$\begin{aligned} \widehat{p}_n(A, B_k, s) - \beta(s) &\geq \widehat{p}_n(A, B_k, X(l+1, m)) - \beta(X(l, m)) \\ (43) \quad &\geq \widehat{p}_n(A, B_k, X(l+1, m)) - \beta(X(l+1, m)) + \frac{b-a}{m}, \end{aligned}$$

$$\begin{aligned} \widehat{p}_n(A, B_k, s) - \beta(s) &\leq \widehat{p}_n(A, B_k, X(l, m)) - \beta(X(l+1, m)) \\ (44) \quad &\leq \widehat{p}_n(A, B_k, X(l, m)) - \beta(X(l, m)) - \frac{b-a}{m}, \end{aligned}$$

because β and $\widehat{p}_n(A, B_k, \cdot)$ are decreasing functions. Let $0 \leq s \leq t$. There are two possibilities: either $s \in [X(m-1, m), t]$, or there exists $0 \leq l \leq m-2$ such that $s \in [X(l, m), X(l+1, m)[$. Together with (43) and (44),

$$\sup_{0 \leq s \leq t} \left| \widehat{p}_n(A, B_k, s) - \beta(s) \right| \leq \alpha_{n,m} + \frac{b-a}{m},$$

where

$$\alpha_{n,m} = \sup_{0 \leq l \leq m} \left| \widehat{p}_n(A, B_k, X(l, m)) - \beta(X(l, m)) \right|.$$

Let

$$\Omega_m = \bigcap_{l=0}^m \Upsilon_{X(l,m)} \quad \text{and} \quad \Omega_\infty = \bigcap_{m \geq 2} \Omega_m.$$

Consequently, when n goes to infinity

$$\forall \omega \in \Omega_m, \quad \alpha_{n,m}(\omega) \rightarrow 0.$$

In addition, $\mathbf{P}_x(\Omega_m) = 1$ because Ω_m is a finite intersection of sets with probability one. Finally,

$$\limsup_{n \rightarrow +\infty} \sup_{0 \leq s \leq t} \left| \widehat{p}_n(A, B_k, s) - \beta(s) \right| \leq \inf_{m \geq 2} \frac{b-a}{m} \quad \mathbf{P}_x\text{-a.s.},$$

since $\mathbf{P}_x(\Omega_\infty) = 1$ as a countable intersection of sets with probability one. \square

Finally, our main result of convergence lies in the following theorem.

Theorem 6.11. *Let \mathcal{K} be a compact subset of E and $\xi \in E$. For any $\epsilon, \eta > 0$, there exist an integer N and a finite partition $P = (A_l)$ of \mathcal{K} such that, for any $0 < t < \min_l t^*(A_l)$, there exists for each couple (l, k) , a sequence $(\beta_n(A_l, \beta_k))_{n \geq 0}$ (depending on t), which almost surely tends to 0, such that for any $n \geq N$, for any $0 < r_1 < r_2 < t$,*

$$\mathbf{P}_\xi \left(\sup_{x \in \mathcal{K}} \sup_{r_1 \leq s \leq r_2} \left| \sum_{k,l} \widehat{l}_{n, \beta_n(A_l, B_k), t}(A_l, B_k, s) \widehat{p}_n(A_l, B_k, s) \mathbf{1}_{\{x \in A_l\}} - f(x, s) \right| > \eta \right) < \epsilon.$$

Proof. Let (A_l) a partition of \mathcal{K} . Let us fix l . Using Proposition 5.13 and Proposition 6.10, there exists a family of sequences $\{(\beta_n(A_l, B_k))_{n \geq 0}\}_{1 \leq k \leq p}$, such that

$$\sup_{r_1 \leq s \leq r_2} \left| \sum_{k=1}^p \widehat{l}_{n, \beta_n(A_l, B_k), t}(A_l, B_k, s) \widehat{p}_n(A_l, B_k, s) - \frac{1}{\nu(A_l)} \sum_{k=1}^p \widetilde{l}(A_l, B_k, s) \int_{A_l \times B_k} H(u, y, s) \nu(du) \mu(dy) \right| \xrightarrow{\mathbf{P}_\xi} 0,$$

when n goes to infinity. In addition, on the strength of Proposition 6.7, for any $x \in \mathcal{K}$, the distance

$$\sup_{r_1 \leq s \leq r_2} \left| f(x, s) - \sum_{l=1}^{|P|} \frac{\mathbf{1}_{\{x \in A_l\}}}{\nu(A_l)} \sum_{k=1}^p \widetilde{l}(A_l, B_k, s) \int_{A_l \times B_k} H(u, y, s) \nu(du) \mu(dy) \right|$$

is arbitrarily small. The triangle inequality immediately ensures the result. \square

In the previous theorem, if the compact subset \mathcal{K} is close to the state space E , then the lower bound of the $t^*(A_k)$'s is small, for any partition (A_k) . Therefore, one estimates the conditional density f on a large part of E , but within a small time interval. Conversely, if \mathcal{K} is chosen centered in E , one may estimate f on a small part of the state space, but within a long time.

7. EXAMPLE

In this section, we present a short simulation study for illustrating the convergence result stated in Theorem 6.11. We consider a PDMP $(X_t)_{t \geq 0}$ defined on the state space $\mathcal{D} \times \mathcal{I}$, where

$$\mathcal{D} = \{x \in \mathbf{R}^2 : \|x\|_2 \leq 1\} \quad \text{and} \quad \mathcal{I} =]0, 2\pi[.$$

In addition, we assume that the process starts from $X_0 = (0, 0, \pi)$. Let us consider $x = (x_1, x_2) \in \mathcal{D}$ and $\theta \in \mathcal{I}$. The flow Φ satisfies,

$$\forall t \geq 0, \quad \Phi((x, \theta), t) = (x_1 + t \cos(\theta), x_2 + t \sin(\theta), \theta).$$

The process $(X_t)_{t \geq 0}$ has two components: intuitively, the first one represents the location in \mathcal{D} , while the second one models the direction of the deterministic motion which takes place in \mathcal{D} . The jump rate λ is given by $\lambda((x, \theta)) = 5 + \|x\|_2$. Since λ does not depend on θ , we will write $\lambda(x)$. Finally, the transition kernel Q is defined for any $A \in \mathcal{B}(\mathbf{R}^2)$ and $B \in \mathcal{B}(\mathbf{R})$, by

$$Q((x, \theta), A \times B) = \frac{1}{K_x} \int_A \mathbf{1}_{\mathcal{D}}(y) \exp\left(-\frac{1}{2\sigma^2} \|y - x\|_2^2\right) dy \int_B \mathbf{1}_{\mathcal{I}}(u) du,$$

with K_x as the normalizing constant and σ^2 as a parameter of variance. When a jump occurs, the new location on \mathcal{D} is chosen according to a gaussian distribution centered in the previous position. The new angle is randomly chosen in \mathcal{I} .

This process may model the movement of a bacteria in a closed environment (see for instance [14]). The bacteria moves on a line with a constant speed. It spontaneously and randomly changes its direction. During the rotation, the location may be a little modified (the parameter σ^2 must be chosen small, $\sigma^2 = 10^{-4}$ in our simulation study). Next, the bacteria moves again on a line. A change of direction occurs also when the bacteria tries to run through the boundary of its environment. In our model, the jump rate λ depends only on the distance between the bacteria and the origin.

In the sequel, we focus our attention on the estimation of the conditional density $f(x_0, t)$ for $x_0 = (0, 0)$. We give an explicit formula of $f(x_0, t)$,

$$\forall t \geq 0, f(x_0, t) = (5 + t) \exp(-t(5 + t/2)).$$

For any $y \in E$ and t , we choose to approximate the jump rate $\tilde{\lambda}(x_0, y, t)$ by $\tilde{l}(A, B_k, t)$, with $A =]-\varepsilon, \varepsilon[$, $\varepsilon = 0.1$, and $B_1 = A$, $B_2 = \mathcal{D} \setminus A$, that is, (B_k) is the simplest partition of \mathcal{D} that one may consider. In this context, $t^*(A) = 0.9$. Therefore, we decide to estimate $\tilde{l}(A, B_k, t)$ by integrating between 0 and 0.8, which is less than $t^*(A)$. Finally, we provide an estimate of f within the interval $[0.05, 0.75]$, which is a proper subset of $[0, 0.8]$.

We simulate a long trajectory of the process: the observation of 50000 jumps is available to estimate f . The number of visits in A is 5330. In addition, the chosen bandwidth $\beta_n(A, B_k)$ can be written in the following way,

$$\beta_n(A, B_k) = \frac{1}{h_n(A, B_k)^\alpha},$$

where $h_n(A, B_k)$ denotes the random number of visits in A followed by a visit in B_k , and $\alpha = 1/3$. Figure 1 is given to illustrate the good behavior of the estimator of f .

APPENDIX A. PROOF OF PROPOSITION 5.8

For the first conditional independence, let h_1, \dots, h_n be some bounded measurable functions mapping from \mathbf{R}_+ into \mathbf{R} . We have $\mathcal{G}_n \subset \mathcal{G}_n \vee \sigma(\delta_{n-1})$ and, by (1), $h_n(S_n)$ is $\mathcal{G}_n \vee \sigma(\delta_{n-1})$ -measurable. Thus,

$$\begin{aligned} & \mathbf{E}_{\nu_0}[h_1(S_1) \dots h_n(S_n) | \mathcal{G}_n] \\ &= \mathbf{E}_{\nu_0}[\mathbf{E}_{\nu_0}[h_1(S_1) \dots h_n(S_n) | \sigma(\delta_{n-1}) \vee \mathcal{G}_n] | \tilde{\mathcal{G}}_n] \\ (45) \quad &= \mathbf{E}_{\nu_0}[h_n(S_n) \mathbf{E}_{\nu_0}[h_1(S_1) \dots h_{n-1}(S_{n-1}) | \sigma(\delta_{n-1}) \vee \mathcal{G}_n] | \mathcal{G}_n]. \end{aligned}$$

Furthermore, always by (1), $\sigma(\delta_{n-1}) \vee \mathcal{G}_n \subset \sigma(\delta_{n-1}) \vee \mathcal{G}_{n-1} \vee \sigma(\varepsilon_{n-1})$. Consequently,

$$\begin{aligned} & \mathbf{E}_{\nu_0}[h_1(S_1) \dots h_{n-1}(S_{n-1}) | \sigma(\delta_{n-1}) \vee \mathcal{G}_n] \\ (46) \quad &= \mathbf{E}_{\nu_0}[\mathbf{E}_{\nu_0}[h_1(S_1) \dots h_{n-1}(S_{n-1}) | \sigma(\delta_{n-1}) \vee \mathcal{G}_{n-1} \vee \sigma(\varepsilon_{n-1})] | \sigma(\delta_{n-1}) \vee \mathcal{G}_n]. \end{aligned}$$

Nevertheless, $h_1(S_1) \dots h_{n-1}(S_{n-1})$ is $\mathcal{G}_{n-1} \vee \sigma(\delta_0, \dots, \delta_{n-2})$ -measurable, and moreover,

$$\sigma(\varepsilon_{n-1}) \vee \sigma(\delta_{n-1}) \perp \mathcal{G}_{n-1} \vee \sigma(\delta_0, \dots, \delta_{n-2}).$$

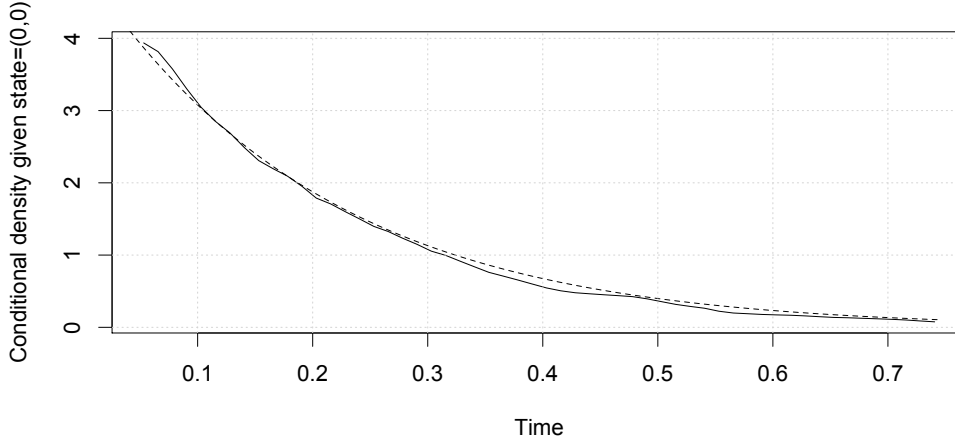


FIGURE 1. Estimation of the conditional density $f(x_0, t)$ with $x_0 = (0, 0)$ and $0.05 \leq t \leq 0.75$. The estimate is drawn in solid line, while the exact density is in dashed line.

Together with (3) [9, page 308],

$$\mathbf{E}_{\nu_0}[h_1(S_1) \dots h_{n-1}(S_{n-1}) | \sigma(\delta_{n-1}) \vee \mathcal{G}_{n-1} \vee \sigma(\varepsilon_{n-1})] = \mathbf{E}_{\nu_0}[h_1(S_1) \dots h_{n-1}(S_{n-1}) | \mathcal{G}_{n-1}].$$

Finally, with (46),

$$\mathbf{E}_{\nu_0}[h_1(S_1) \dots h_{n-1}(S_{n-1}) | \sigma(\delta_{n-1}) \vee \mathcal{G}_n] = \mathbf{E}_{\nu_0}[h_1(S_1) \dots h_{n-1}(S_{n-1}) | \mathcal{G}_{n-1}].$$

Thus, by (45),

$$\mathbf{E}_{\nu_0}[h_1(S_1) \dots h_n(S_n) | \mathcal{G}_n] = \mathbf{E}_{\nu_0}[h_1(S_1) \dots h_{n-1}(S_{n-1}) | \tilde{\mathcal{G}}_{n-1}] \mathbf{E}_{\nu_0}[h_n(S_n) | \mathcal{G}_n].$$

Therefore, by a straightforward induction, we have

$$\mathbf{E}_{\nu_0}[h_1(S_1) \dots h_n(S_n) | \mathcal{G}_n] = \prod_{i=1}^n \mathbf{E}_{\nu_0}[h_i(S_i) | \mathcal{G}_i].$$

Taking for $j \neq i$, $h_j = \mathbf{1}$, leads to

$$(47) \quad \mathbf{E}_{\nu_0}[h_i(S_i) | \mathcal{G}_n] = \mathbf{E}_{\nu_0}[h_i(S_i) | \mathcal{G}_i].$$

Hence,

$$\mathbf{E}_{\nu_0}[h_1(S_1) \dots h_n(S_n) | \mathcal{G}_n] = \prod_{i=1}^n \mathbf{E}_{\nu_0}[h_i(S_i) | \mathcal{G}_i].$$

This shows that

$$\bigvee_{j \neq i} \sigma(S_j) \perp_{\mathcal{G}_n} \sigma(S_i),$$

and this directly induces the expected result. For the second conditional independence, let now $h_1 : \mathbf{R}_+ \rightarrow \mathbf{R}$ and $h_2 : E^{n+1} \rightarrow \mathbf{R}$ be some bounded measurable functions.

$\sigma(Z_{i-1}, Z_i) \subset \mathcal{G}_n$, thus,

$$\begin{aligned}
 & \mathbf{E}_{\nu_0} [h_1(S_i)h_2(Z_0, \dots, Z_n) | \sigma(Z_{i-1}, Z_i)] \\
 &= \mathbf{E}_{\nu_0} \left[\mathbf{E}_{\nu_0} [h_1(S_i)h_2(Z_0, \dots, Z_n) | \mathcal{G}_n] | \sigma(Z_{i-1}, Z_i) \right] \\
 (48) \quad &= \mathbf{E}_{\nu_0} \left[h_2(Z_0, \dots, Z_n) \mathbf{E}_{\nu_0} [h_1(S_i) | \mathcal{G}_n] | \sigma(Z_{i-1}, Z_i) \right].
 \end{aligned}$$

We shall prove that

$$(49) \quad \mathbf{E}_{\nu_0} [h_1(S_i) | \mathcal{G}_n] = \mathbf{E}_{\nu_0} [h_1(S_i) | \sigma(Z_{i-1}, Z_i)].$$

By (47), we have

$$\mathbf{E}_{\nu_0} [h_1(S_i) | \mathcal{G}_n] = \mathbf{E}_{\nu_0} [h_1(S_i) | \sigma(Z_0, \dots, Z_i)].$$

Therefore, in order to state (49), we have to prove that

$$(50) \quad \sigma(S_i) \underset{\sigma(Z_{i-1}, Z_i)}{\perp} \sigma(Z_0, \dots, Z_{i-2}).$$

From the dynamic (1), we have

$$\sigma(\delta_{i-1}, \varepsilon_{i-1}) \perp \sigma(Z_0, \dots, Z_{i-2}) \vee \sigma(Z_{i-1}).$$

Thus, in the light of Proposition 6.8 of [17] in the direction \Rightarrow , we have

$$(51) \quad \sigma(\delta_{i-1}, \varepsilon_{i-1}) \underset{\sigma(Z_{i-1})}{\perp} \sigma(Z_0, \dots, Z_{i-2}).$$

Furthermore, it is easy to see that

$$(52) \quad \sigma(Z_0, \dots, Z_{i-2}) \underset{\sigma(Z_{i-1})}{\perp} \sigma(Z_{i-1}).$$

Therefore, with (51), (52) and Proposition 6.8 of [17] again, but in the direction \Leftarrow , we have

$$\sigma(Z_{i-1}, \delta_{i-1}, \varepsilon_{i-1}) \underset{\sigma(Z_{i-1})}{\perp} \sigma(Z_0, \dots, Z_{i-2}).$$

Furthermore, we have the equality of the σ -fields

$$\sigma(Z_{i-1}, Z_i, \delta_{i-1}, \varepsilon_{i-1}) = \sigma(Z_{i-1}, \delta_{i-1}, \varepsilon_{i-1}),$$

since Z_i is generated by Z_{i-1} and the random errors δ_{i-1} and ε_{i-1} . Thus,

$$\sigma(Z_i) \vee \sigma(Z_{i-1}, \delta_{i-1}, \varepsilon_{i-1}) \underset{\sigma(Z_{i-1})}{\perp} \sigma(Z_0, \dots, Z_{i-2}).$$

Thus, on the strength of Proposition 6.8 of [17] again, in the direction \Rightarrow , we have

$$\sigma(Z_{i-1}, \delta_{i-1}, \varepsilon_{i-1}) \underset{\sigma(Z_{i-1}, Z_i)}{\perp} \sigma(Z_0, \dots, Z_{i-2}).$$

This states (50) since the time S_i is generated from Z_{i-1} and the error δ_{i-1} . Therefore, we prove (49). As a consequence, plugging (48) and (49) yields to

$$\begin{aligned}
 & \mathbf{E}_{\nu_0} [h_1(S_i)h_2(Z_0, \dots, Z_n) | \sigma(Z_{i-1}, Z_i)] \\
 &= \mathbf{E}_{\nu_0} [h_1(S_i) | \sigma(Z_{i-1}, Z_i)] \mathbf{E}_{\nu_0} [h_2(Z_0, \dots, Z_n) | \sigma(Z_{i-1}, Z_i)],
 \end{aligned}$$

showing the expected result. \square

REFERENCES

- [1] AALEN, O. O. *Statistical inference for a family of counting processes*. ProQuest LLC, Ann Arbor, MI, 1975. Thesis (Ph.D.)—University of California, Berkeley.
- [2] AALEN, O. O. Weak convergence of stochastic integrals related to counting processes. *Z. Wahrscheinlichkeitstheorie und Verw. Gebiete* 38, 4 (1977), 261–277.
- [3] AALEN, O. O. Nonparametric inference for a family of counting processes. *Ann. Statist.* 6, 4 (1978), 701–726.
- [4] AALEN, O. O., ANDERSEN, P. K., BORGAN, Ø., GILL, R. D., AND KEIDING, N. History of applications of martingales in survival analysis. *J. Électron. Hist. Probab. Stat.* 5, 1 (2009), 28.
- [5] ANDERSEN, P. K., BORGAN, Ø., GILL, R. D., AND KEIDING, N. *Statistical models based on counting processes*. Springer Series in Statistics. Springer-Verlag, New York, 1993.
- [6] AZAÏS, R., DUFOUR, F., AND GÉGOUT-PETIT, A. Nonparametric estimation of the jump rate for non-homogeneous marked renewal processes. *Preprint*.
- [7] BERAN, J. Nonparametric regression with randomly censored survival data, 1981. Technical report, Dept. Statist. Univ. California, Berkeley.
- [8] CHIQUET, J., AND LIMNIOS, N. A method to compute the transition function of a piecewise deterministic Markov process with application to reliability. *Statist. Probab. Lett.* 78, 12 (2008), 1397–1403.
- [9] CHUNG, K. L. *A course in probability theory*, second ed. Academic Press, New York-London, 1974. Probability and Mathematical Statistics, Vol. 21.
- [10] COMTE, F., GAÏFFAS, S., AND GUILLOUX, A. Adaptive estimation of the conditional intensity of marker-dependent counting processes. *Ann. Inst. H. Poincaré Probab. Statist.* 47, 4 (2011), 1171–1196.
- [11] COX, D. R. Regression models and life-tables. *J. Roy. Statist. Soc. Ser. B* 34 (1972), 187–220.
- [12] DABROWSKA, D. M. Nonparametric regression with censored survival time data. *Scand. J. Statist.* 14, 3 (1987), 181–197.
- [13] DAVIS, M. H. A. *Markov models and optimization*, vol. 49 of *Monographs on Statistics and Applied Probability*. Chapman & Hall, London, 1993.
- [14] FONTBONA, J., GUÉRIN, H., AND MALRIEU, F. Quantitative estimates for the long time behavior of a PDMP describing the movement of bacteria. *Preprint*.
- [15] HERNÁNDEZ-LERMA, O., AND LASSERRE, J. B. *Markov chains and invariant probabilities*, vol. 211 of *Progress in Mathematics*. Birkhäuser Verlag, Basel, 2003.
- [16] JACOBSEN, M. *Statistical analysis of counting processes*, vol. 12 of *Lecture Notes in Statistics*. Springer-Verlag, New York, 1982.
- [17] KALLENBERG, O. *Foundations of modern probability*, second ed. Probability and its Applications (New York). Springer-Verlag, New York, 2002.
- [18] LI, G., AND DOSS, H. An approach to nonparametric regression for life history data using local linear fitting. *Ann. Statist.* 23, 3 (1995), 787–823.
- [19] MARTINUSSEN, T., AND SCHEIKE, T. H. *Dynamic regression models for survival data*. Statistics for Biology and Health. Springer, New York, 2006.
- [20] MCKEAGUE, I. W., AND UTIKAL, K. J. Inference for a nonlinear counting process regression model. *Ann. Statist.* 18, 3 (1990), 1172–1187.
- [21] MEYN, S., AND TWEEDIE, R. L. *Markov chains and stochastic stability*, second ed. Cambridge University Press, Cambridge, 2009.
- [22] OUVRARD, J.-Y. *Probabilités 2*. Deuxième édition, Master, Agrégation. Cassini, 2004.
- [23] RAMLAU-HANSEN, H. Smoothing counting process intensities by means of kernel functions. *Ann. Statist.* 11, 2 (1983), 453–466.
- [24] STUTE, W. Conditional empirical processes. *Ann. Statist.* 14, 2 (1986), 638–647.
- [25] UTIKAL, K. J. Nonparametric inference for a doubly stochastic Poisson process. *Stochastic Process. Appl.* 45, 2 (1993), 331–349.
- [26] UTIKAL, K. J. Nonparametric inference for Markovian interval processes. *Stochastic Process. Appl.* 67, 1 (1997), 1–23.

E-mail address: romain.azais@inria.fr

E-mail address: dufour@math.u-bordeaux1.fr

E-mail address: anne.petit@u-bordeaux2.fr

INRIA BORDEAUX SUD-OUEST, TEAM CQFD, FRANCE AND UNIVERSITÉ BORDEAUX, IMB, CNRS UMR 5251, 351, COURS DE LA LIBÉRATION, 33405 TALENCE CEDEX, FRANCE.